



# Case Study: The California Digital Library

---

2<sup>nd</sup> Workshop on the Open Archives  
Initiative: Gaining Independence  
with e-prints archives and OAI

October 2002, CERN

John Ober, Director of Education and  
Strategic Innovation

California Digital Library  
[www.cdlib.org](http://www.cdlib.org)



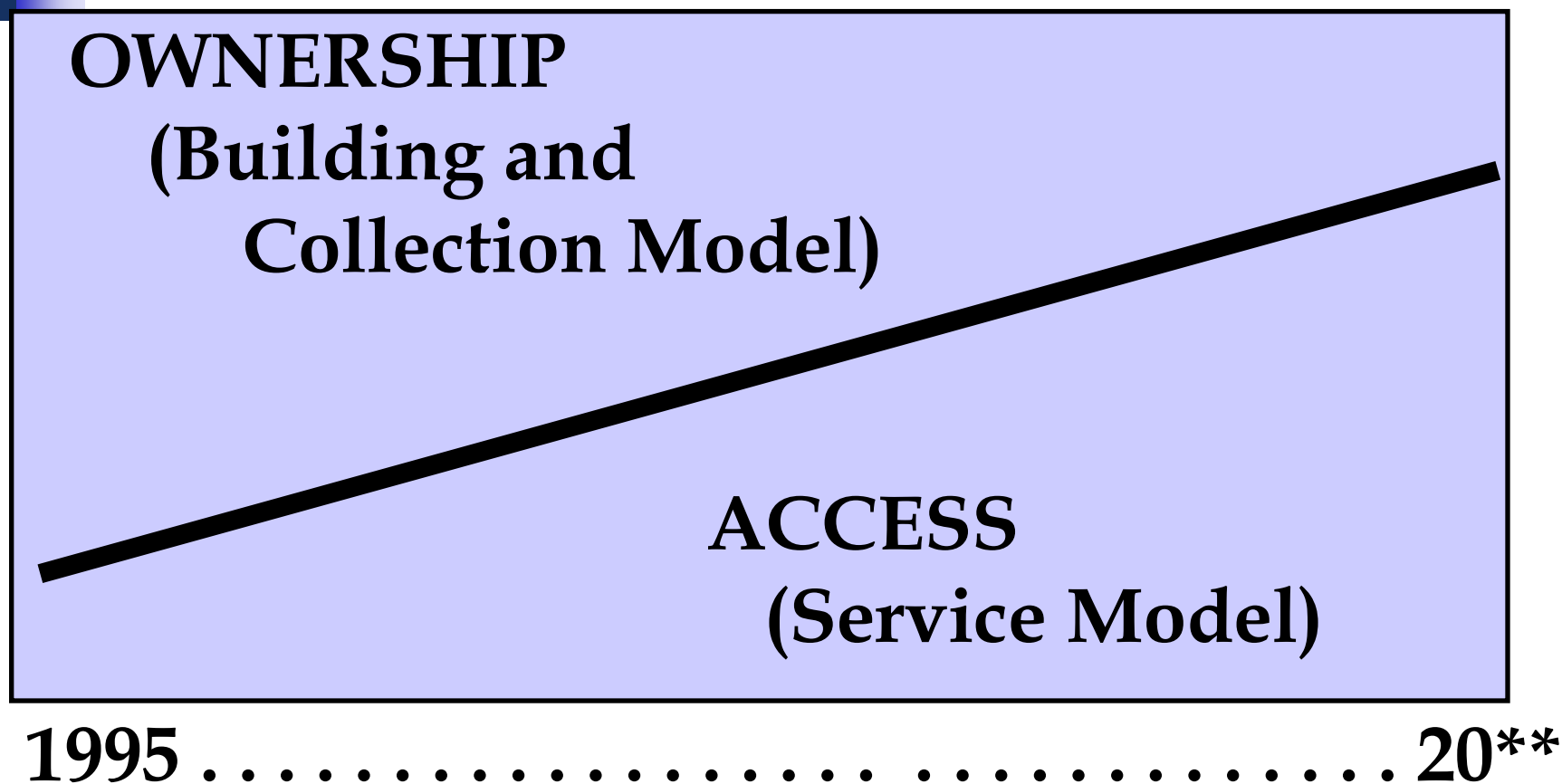
# CDL History

---

- Founded in 1997; digital “co-library” of the 10 Univ. of California campuses
- Response to *crisis* in scholarly communication and the *opportunity* presented by digital technologies and the Web
- Charged to create a **comprehensive system for the management** of digital scholarly information
- “...and facilitating and supporting scholar-led innovations in scholarly communication.” (CDL mission statement)



## Challenges, opportunities





## Challenges, opportunities

**Traditional Forms of  
Scholarly Communication**

**Alternative Forms of  
Scholarly Communication**

**2000 . . . . . 20\*\***



## CDL Overview

---

- \$14 million budget
- 50 employees on site; ~10 on campuses
- Digital resources
  - 200 databases licensed
  - ~6 locally-hosted databases incl. 24 million record catalog
  - 7000 online journals licensed
  - ~200,000 images and documents hosted
  - ~ 2500 social science data sets
- ~25 million searches/year on hosted data



## CDL's eScholarship Program was started to:

---

- Monitor emerging technologies and experimental applications
  - Use of repositories in physics, economics, etc.
  - XML/structured text, etc.
- Understand needs of scholars and researchers
  - Readiness to experiment, desire to enhance communication, to lower barriers to access
  - Concern with Peer Review and Permanence of the scholarly record

**Conclusion:** Experiment w/components of scholarly communication



## eScholarship: *Use of technologies*

---

- *Traditional: familiar things in better way*
  - Digitally reformatted Books
  - Digital Journals
  - New digital monographs
- *Experimental: newer things...*
  - Discipline-based Digital Repositories for self-submission and self-archiving
  - Standards-based metadata and data sets to ensure interoperability
  - “Publication” of [Dynamic] Digital Data Sets



## eScholarship: *Discipline-based experiments*

---

- Repository of Working Papers in Social Sciences and Humanities
- [California] International and Area Studies: Digital pub of monographs from repository articles; new business model
- Dermatology Online Journal and eNvironment: enhancements to journal articles, supplementary material
- Optical Society of America (w/ Univ of Cal Press) multi-part digital and print reference work



**Scholarship** Scholar-led innovations in scholarly communication **CDL**

Search Feedback About Us Site Index

**CONQUERING WOMEN**  
THE FIRST  
WORLD WAR  
THE GERMAN  
CENTRAL  
POWER

International & Area  
Studies Publications

Search this Web Site:  
 **go >**

**eScholarship  
Communities**

View Publishing Models

**Dermatology**

**Electronic Cultural  
Atlas Initiative**

**Environmental  
Studies**

**International and  
Area Studies**

**Tobacco Control**

an initiative of the California Digital Library



© Copyright 2002 The Regents of the University of California  
http://escholarship.cdlib.org/ · Site management: (510) 987-0476  
eScholarship, California Digital Library, 415 20th Street, 4th Floor, Oakland, CA 94612-2901



## eScholarship: *Collaborations & Partnerships*

---

- University of California Press: leverage their expertise (editing, business models)
- SPARC: grant supports experiment to use adv tech to communicate in new ways
- bepress: toolkit for peer review journals; co-developing tools for digital repository



## bepress

---

- UCB faculty start-up company
- EdiKit: Web-based peer review journal software.  
Features include:
  - Easy article submission process
  - Assignment and tracking of reviewers
  - Automatic email notifications
  - Pre-publication review process
  - Publication of single articles or entire issue



## CDL & bepress

---

- CDL has two relationships with bepress:
  - license for their EdiKit software for e-journal, e-book, and working paper repository publication and management
  - A co-development agreement that enables CDL to specify future enhancements
- CDL positioned to take a commercial-grade product and tailor it to library needs (e.g., interoperability with other repositories using OAI)



# UNIVERSITY OF CALIFORNIA eScholarship Repository



## eScholarship

### **institutes**

### **search**

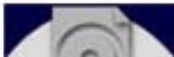
### **help**

### **my account**

### **search melvyl**

The eScholarship Repository offers faculty a central location for depositing pre-publication scholarship. The repository, sponsored by the [California Digital Library](#) of the University of California, provides persistent access to working papers and makes them easily discoverable. The CDL's [eScholarship initiative](#), whose mission is to facilitate and support scholar-led innovations in scholarly communication, is providing this and other services in response to an expressed need for alternative publishing mechanisms. For more information, please see the repository [press release](#) (in Adobe Acrobat format).

Search all papers:

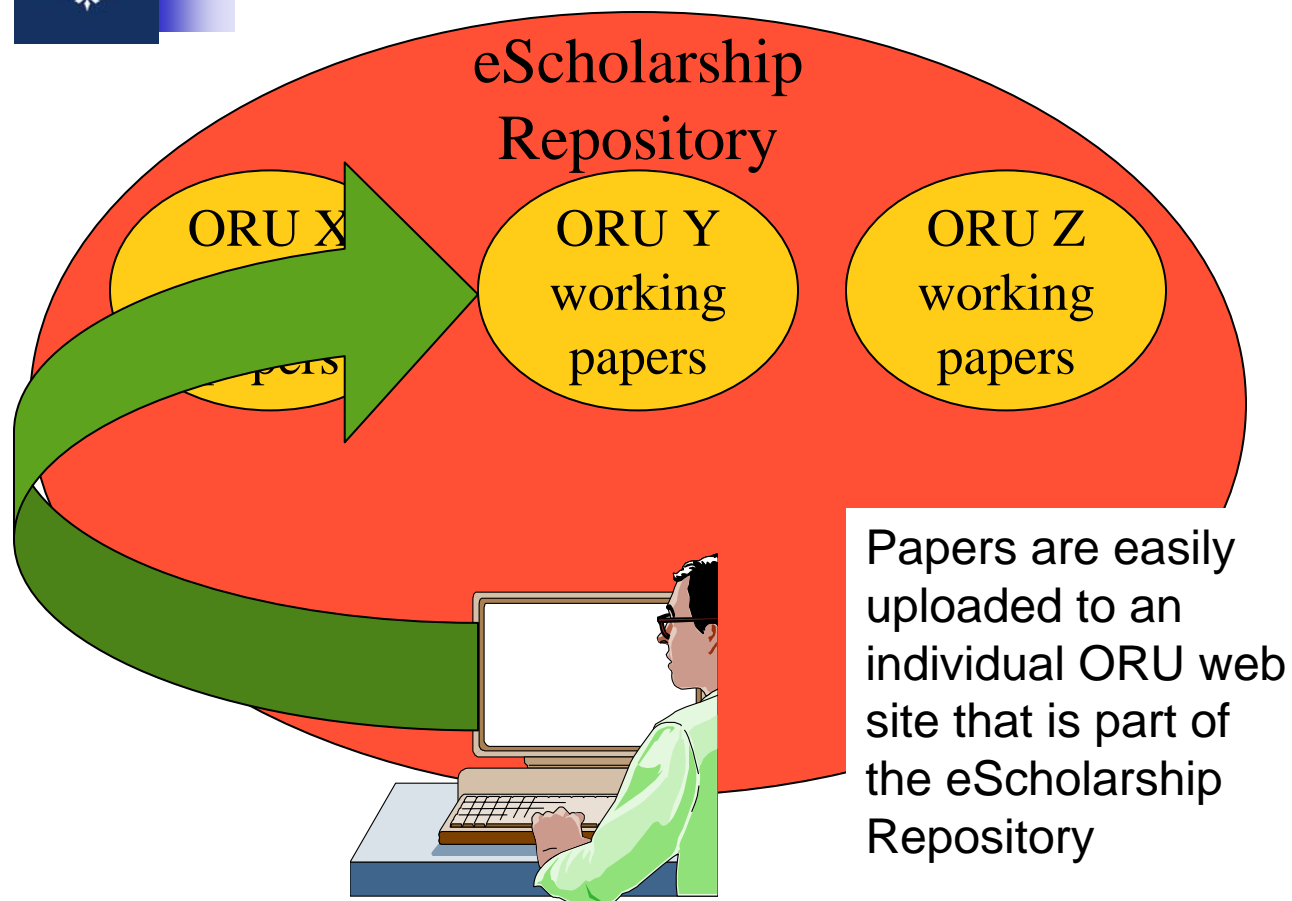


## Working Paper Sites

- [Berkeley Olin Program in Law & Economics, Working Paper Series](#)
- [Institute of Business and Economic Research](#)
- [Institute of Industrial Relations](#)
- [Institute of Transportation Studies](#)
- [University of California Transportation Center](#)

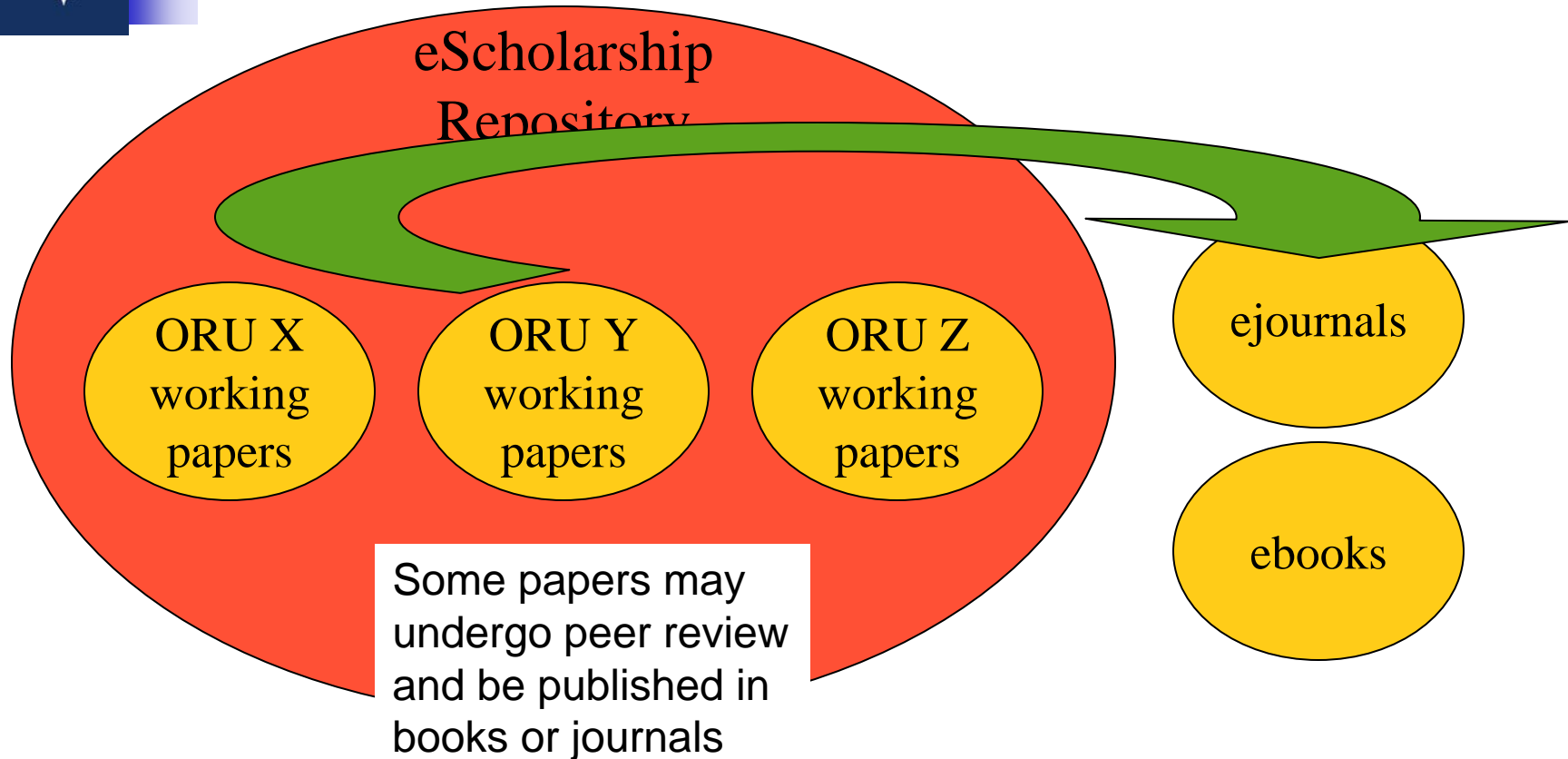


# Repositories: Faculty Perspective



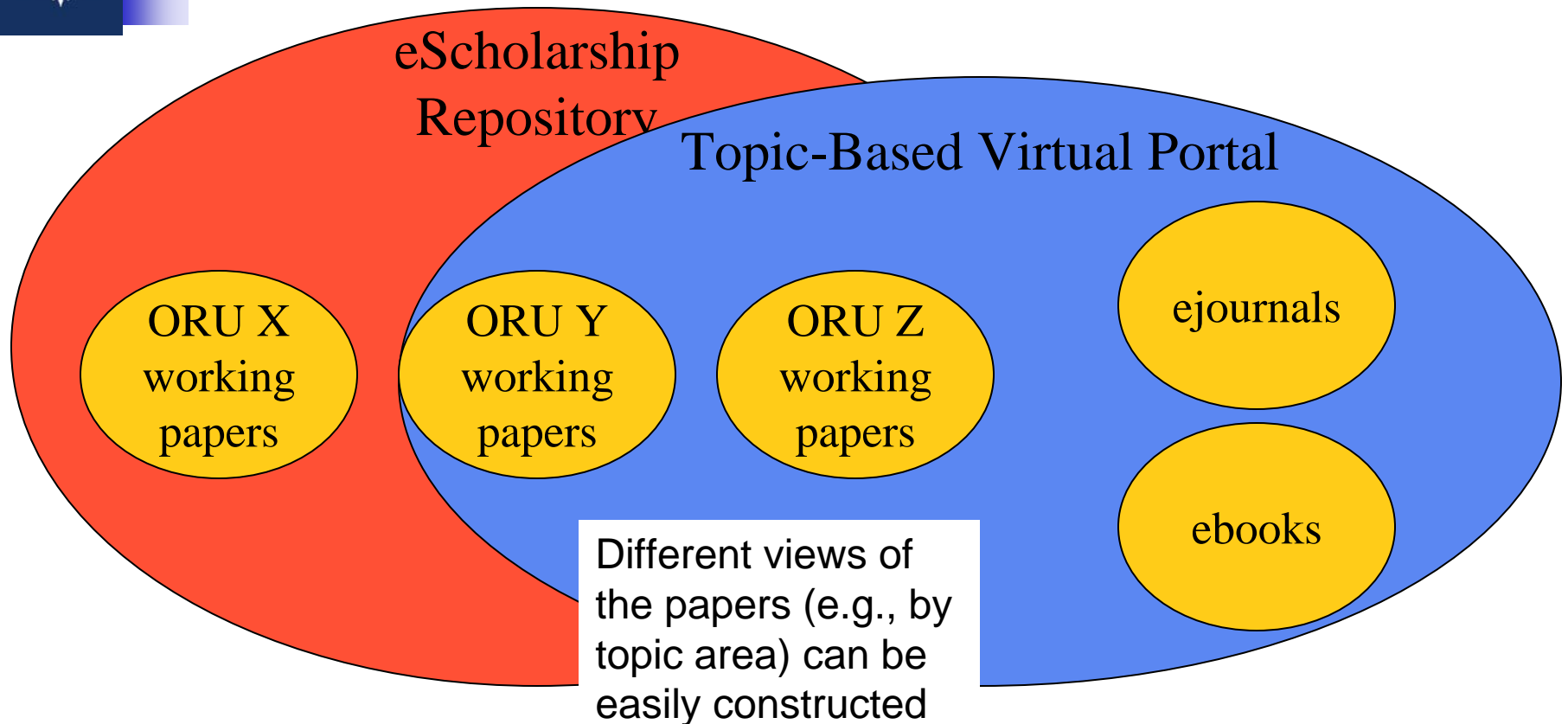


# Repositories: Faculty Perspective





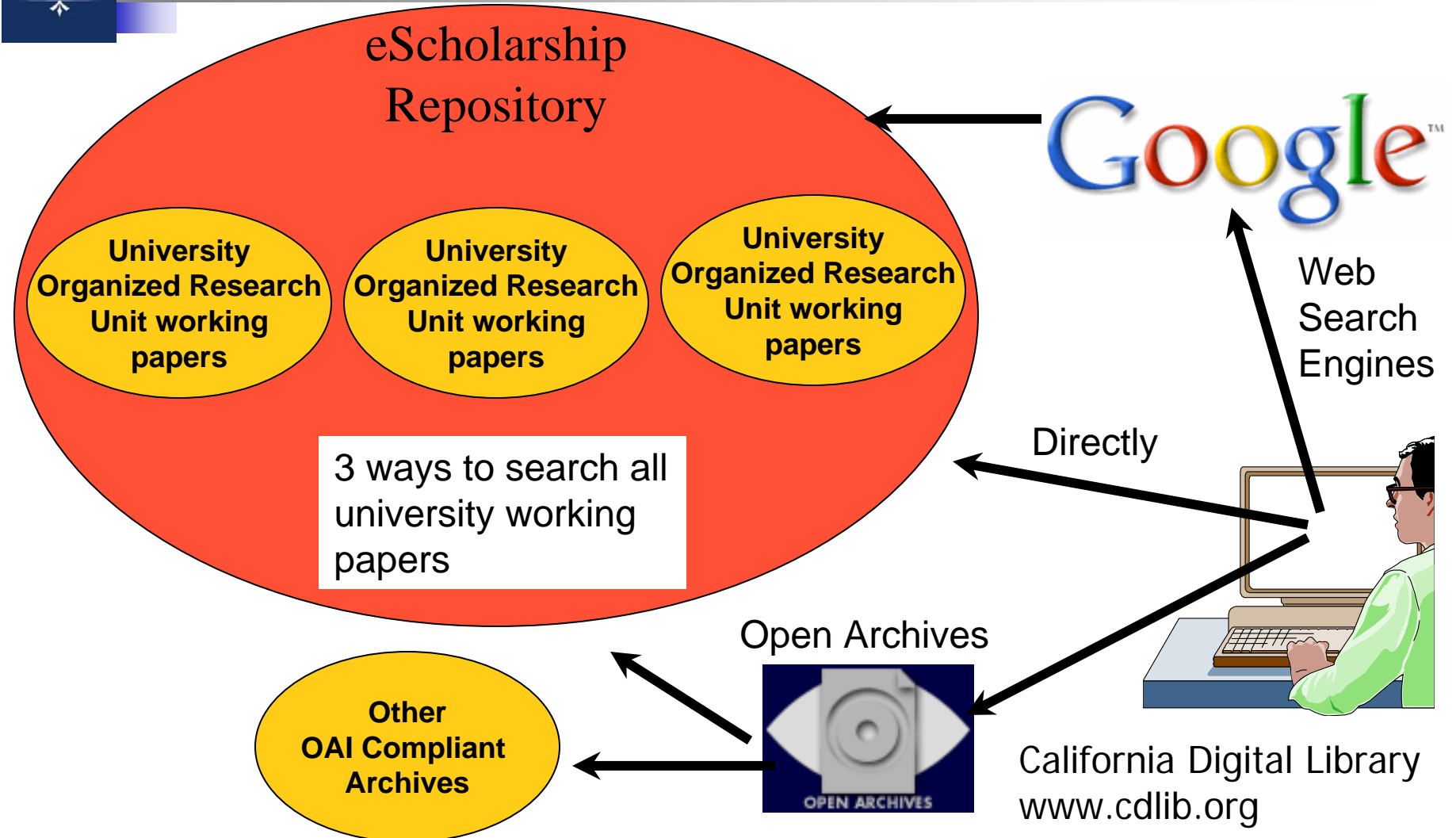
# Repositories: User Perspective

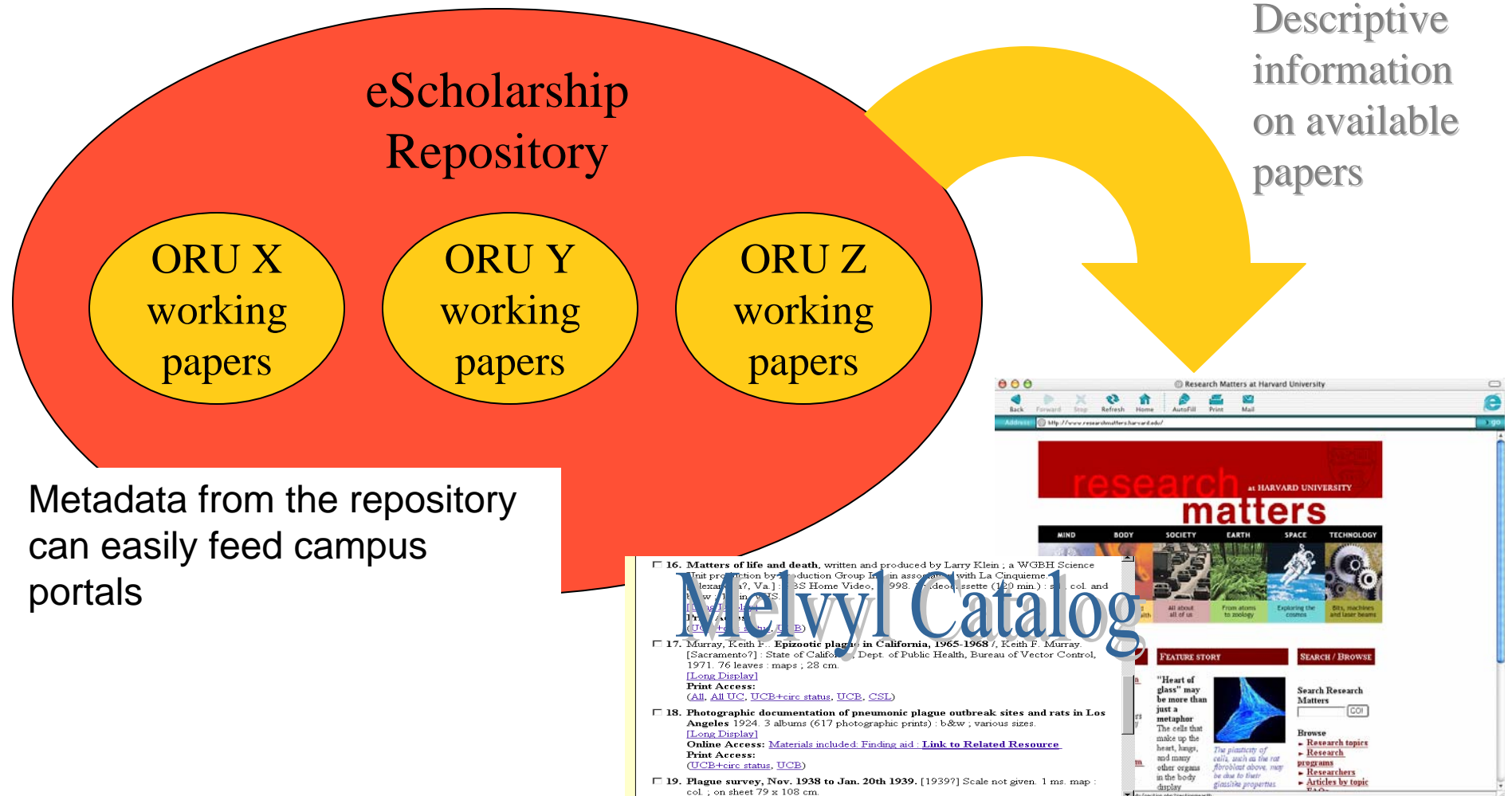






# Repositories: User Perspective







## eScholarship Repository

---

- Federation across arbitrary subsets as well as branding by source
- CDL's commitment to maintain over time (cf: CDL's preservation efforts)
- Formats can be normalized -> structured text
- Lowers entry barriers for digital publication
- Competition for commercial pre-print ventures with pay to deposit and pay to subscribe



## Putting pieces together

---

- EdiKit: working paper repositories; journal and anthology publication; “data rescue” (from faculty websites)
- CDL’s XML/METS based content management system: book publication; comprehensive object searching; preservation

Plus

- Open archives initiative

Equals

- Flexibly managed content ready for inclusion in wider services; ready to seed UC-built services



## CDL's Use of OAI

---

- All eScholarship objects are available through OAI harvesting
- All non-restricted CDL objects are (or soon will be) available through OAI harvesting (e.g. EADs, digitized objects)
- CDL uses OAI to harvest metadata from collection X for inclusion in a CDL access service across collections
- CDL may build OAI-based services (e.g. Social Science scholarship repository)
- CDL is helping/encouraging UC campuses to become data providers and service providers (e.g. UCLA music; UCSF tobacco control)



# Organization and policies for collaboration, innovation

---

- **Funds** – *sought commitment from President, aggressive grant proposal writing, co-investment formulas*
- **Staff** - *invent jobs, aggressive staff development and skills inventorying, borrow staff from partners*
- **Understanding faculty** – *~45 focus groups; senior staff spend ~50%+ of their time communicating/visiting*
- **Technology** - *standards AND leaps of faith, formal advice, open-source, outsourcing*
- **Organization and procedures** - *team/matrix structure; invent policy*



## Surface challenges – One organization innovating for scholarly communication

---

- Acquire expertise
- Assess user needs, declare priority actions
- Design projects and workflows
- Acquire technology
- Fund development



## Deeper challenges

---

- Acquire expertise Maintain and refresh expertise
- Assess user needs, declare priority actions Become scholar-centered (rather than content or technology centered), refresh and revise systems
- Design projects and workflows Design an adaptive, learning organization
- Acquire technology Build or buy technology?; transfer technology, engage in research?
- Fund development Fund collaboration (high overhead), fund contextual activities (public relations; OAI)





## Surface challenges – goals for scholarly communication

---

- Show scholars the benefits in joining
- Build data provider and service provider software
- Build ancillary services (citation referencing, ranking, etc.)
- Increase the number of participants
- Convince administrators to fund infrastructure



## Deeper challenges

---

- Show scholars the benefits in joining/ Find/cultivate scholars who lead the innovation
- Build data provider and service provider software/Integrate across information and access silos
- Build ancillary services (citation referencing, ranking, etc.)/Influence reward system
- Increase the number of participants /Address scale, scope and interoperability issues
- Convince administrators to fund infrastructure /Test and establish viable business models



I'm finished